

Non-Logographic Chinese and the Non-Alphabetic Alphabet

Author(s): Mark Hansell

Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society (1989), pp. 102-113

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

NON-LOGOGRAPHIC CHINESE
AND THE
NON-ALPHABETIC ALPHABET

Mark Hansell
University of California, Berkeley
(1)

0. INTRODUCTION

The classification of writing systems in general, and of the Chinese writing system in particular, has been the cause of much controversy. This paper will examine some of the classificatory difficulties involving the writing systems of Chinese and English. The development of Mandarin loanwords from English, in which Chinese characters can be used phonographically and Roman letters non-alphabetically, suggests that writers/readers have a variety of strategies available to them (including the logographic and phonographic principles). These strategies are options that can be chosen to fit the situation or even used simultaneously, in order to preserve what Wang (1981) calls "optimality" in the writing system, the prototypical relationship between the structure of written and spoken language. The paper will be organized as follows:

1. Classification of writing systems
2. Types of lexical borrowing
3. Chinese borrowing from English
4. Discussion
5. Conclusion

1. CLASSIFICATION OF WRITING SYSTEMS

1.1. Traditional "idealized" view

In most studies of writing, scripts are assumed by definition to be graphic means of rendering spoken language. Scripts are classified according to the level of linguistic structure that corresponds to the smallest independent graphic unit. If a grapheme corresponds to a segment, the script is alphabetic; if it corresponds to a syllable, it is syllabic; if it corresponds to a word, it is logographic. Chinese is considered the logographic script par excellence, and English is considered to be an example (though perhaps not the best) of a language using an alphabetic system. (See Gelb 1963, Pulgram 1976 for detailed discussions of such classifications.)

An idealized logographic script consists of

graphemes that are associated directly to lexical items of the language, with no phonological elements. Phonological and morphological information are supplied by the lexicon of the spoken language as part of the writer/reader's linguistic competence. Since it has no direct connection to pronunciation, an ideal logogram can represent the same word despite varying pronunciations, even across different languages, as long as the identity of the lexical item is preserved. In the traditional view of Chinese as an ideal logographic script, each character represents a single monosyllabic word. These words are cognate across all Chinese dialects, meaning that mutually unintelligible dialects of Chinese are united by a common writing system.

In an ideal alphabetic system, graphemes are directly associated to segments in the phonology of the spoken language. Strings of letters are arranged in a linear order that corresponds to the temporal order of segments in a spoken word, and from the phonological information given by the writing, morphological and lexical information is supplied by the linguistic knowledge of the reader, in the same way that sequences of segments in speech are interpreted. The essential differences between these two types of writing can be summarized as follows: 1) Level of structure linked to graph: lexical item in one case, phoneme in the other. 2) Direct connection to meaningful unit: present in logographic, absent in alphabetic. 3) Direct connection to phonological representation: present in alphabetic, absent in logographic.

1.2. Complications

Applying these ideal categories to actual writing systems leads to many complications. One is that even cursory examination shows that no form of Modern Chinese has monosyllabic words, especially not Mandarin, the form with which we will concern ourselves. The elements linked to graphemes in Chinese are morphemes, which are generally (with interesting exceptions (2)) monosyllabic. Some scholars have suggested that the Chinese writing system should more properly be categorized as morphographic or morphosyllabic. I wholeheartedly agree with this suggestion, and from here on will use the term morphographic to refer to the reality of Chinese writing, reserving logographic to refer to the idealized notion introduced above. This deviation from the logographic ideal changes the level of structure from word to morpheme, however it has no bearing on the other essential differences between logographic and alphabetic writing: a character still represents a meaningful unit, without reference to sound.

A much more fundamental complication lies in the

structure of Chinese characters themselves: by one count 97% of them contain a phonetic element that gives useful hints as to the character's pronunciation (DeFrancis 1984). Though these phonetic elements may be explained away as relics of the historical character-formation process, evidence suggests that they play an important role in reading and writing.

The behavior of readers and writers of English does more to complicate the situation. A glance at the English orthography is enough to destroy any notion that letters are related to phonemes in any regular way (in fact, they are related in many irregular ways). All the evidence indicates that readers and writers make use of spelling irregularities to differentiate words by graphic shape (that is, logographically) in pairs like Sue and Sioux, bee and be (Bolinger 1946) (3). Though these spelling irregularities, like the phonetic elements in the Chinese characters, may be attributed to historical accident in the course of lexical accretion, the fact that they disrupt the purity of the classification does nothing to dissuade native readers and writers from using them. Users of the written language use any available strategy to make their task easier, including a "phonographic" strategy in Chinese and a logographic strategy in English. (4)

In light of the above complications, the differences between the English and Chinese writing systems in terms of the three criteria above become less clear. The level of structure linked to a graph is still different-- a syllable-morpheme for Chinese, and a phone for English-- but connection of written form to meaningful unit is no longer exclusively the province of Chinese (cf. Sue and Sioux), and representation of phonological information in the script is not restricted to languages like English (cf. the phonetic element 皇 in the characters 惶 huang2 "afraid", 煌 huang2 "brilliant", 蝗 huang2 "locust", 鯨 huang2 "sturgeon" etc. (DeFrancis 1984:103)). The distinction comes closest to collapsing in the area of Chinese transliteration of loanwords. French (1976:109) states that "...in one field, namely its representation of loanwords, the Chinese writing system may justly be called phonographic." If this is the case, as many scholars agree, then the obvious question is why would a writing system completely change its character simply because the word it is transcribing is of foreign origin? Is the functioning of a writing system determined by the type of script, by the language it is applied to, or by the relation between the two?

In what follows I will examine the way lexical borrowing from English interacts with writing in the Chinese of contemporary Taiwan. In addition to

"phonographic" uses of Chinese characters, I will also examine the non-alphabetic use of the Roman alphabet, which is a recent and rapidly expanding phenomenon in Taiwan. Before introducing the data, a brief introduction of a theory of lexical borrowing is necessary.

2. LEXICAL BORROWING

I have elsewhere paraphrased Haugen (1950:212) and defined lexical borrowing as "the attempted reproduction in one language of signs previously found in another" (Hansell 1986:10). The sign in question is a Saussurean sign consisting of the linkage of a signified and a signifier (deSaussure 1915). Borrowing, as opposed to code-switching, requires that the signifier be reproduced out of "native" linguistic material, that is, using elements already extant in the L2 system. This is done through a two-step process described by Weinreich (1954): identification and substitution. Identification is the selection of the elements in the L2 system that are most similar to elements of the same level in the L1 system. Since linguistic units like phonemes, morphemes, words etc. are abstract entities that become meaningless outside of the system they belong to, similarity across languages must be measured in terms of qualities that exist outside the system of a particular language: phonetic similarity (in terms of acoustic shape) and similarity in meaning (in terms of similarity of scope of real-world referents). Substitution then substitutes the identified L2 elements for the L1 elements to create the new L2 signifier. For example, on the basis of phonetic similarity, English [t^h] is identified with Japanese [t]; [ow] with [o]; [f] with [f]; and [u] with [u] respectively. Given the Japanese signifier tofu [to^hfu], the identified English elements are substituted to form [t^howfu].

When identification is by phonetic similarity as above, substitution of L2 phones results in a phonetic loan. Identification is also possible in terms of meaning, with substitution at the morphemic or lexical level. The result is a loan-translation, or calque. For example, English hot dog is the model for Mandarin 热狗 re4 gou3 'hot'+ 'dog'. The morphemes hot and re4 are identified through similarity of reference to the quality "hot", and dog and gou3 on the basis of reference to the same animal "dog". It is the combination of these morphemes that is novel in Mandarin, and therein lies the reproduction of the signifier.

One other type of loan is the graphic loan, where a written signifier is vacuously "reproduced". This happens between two languages that use the same script, when the written form of a word is simply copied into L2

and pronounced according to L2 reading rules. For example, we have borrowed the written form of Paris from the French, but have supplied our own pronunciation [p^hʔi:is]. If we had substituted English phones and spelled it accordingly, we should have something like [p^hʔi:ij] spelled Parree. Graphic loans are very common between Chinese and Japanese, and between Chinese dialects, for instance Chinese 科学 ke1 xue2 "science" from Japanese 科学 kagaku. (5)

3. MANDARIN WRITTEN FORMS OF BORROWINGS FROM ENGLISH

The data that follow were collected in Taiwan in 1987. Some of the items were borrowed prior to 1949 and are therefore similar to forms used in Mainland China, others are more recent innovations.

3.1. Transliteration of phonetic loans

In the transliteration of phonetic loans, Chinese characters may be used strictly for their pronunciation value, without their meaning being taken into account at all. For instance:

English	Mandarin form	Literal gloss
#1) MICROPHONE	麥 克 風 mai4 ke4 feng1	'wheat'+ 'conquer'+ 'wind'
#2) DACRON	達 克 龍 da2 ke4 long2	'attain'+ 'conquer'+ 'dragon'
#3) PUDDING	布 丁 bu4 ding1	'cloth'+ 'population'
#4) BASS (GUITAR)	貝 斯 bei4 si1	'cowrie'+ 'this'

The meanings normally associated with the characters in #1-#4 are totally unrelated to their use in these transliterations. The characters here were chosen strictly because their normal reading pronunciations in Mandarin approximate the strings of phones substituted for the English phones.

Although there is a set of characters that appears often in transliterations (克 ke4 in #1 and #2 is one, so is 斯 si1 in #4), their preponderance is only a statistical tendency, transliteration characters are by no means restricted to a finite set. There is no syllabary from which all transliteration characters are chosen (6). Conversely, writers are not wholly free to substitute homophonous characters in transliterations: for instance, 布丁 bu4 ding1 "pudding" could not be written 步叮 or 部釘, even though the pronunciation of all three sequences is identical. Though there is

always variation in new borrowings, a standard is quickly established for any given loanword.

Of course the meaning of a character is not ignored if it is possible to somehow connect it to the meaning of the English model word. In some cases, skillful choice of transliteration characters can suggest something about the meaning of the borrowed term, optionally introducing a degree of semanticity (without changing its status as a phonetic loan). For example:

- #5) JELLY BEAN 結粒(糖) 'coagulate'+ 'pellet'
jie2 li4 (tang2) +('candy')
- #6) The BEATLES 披頭四 'disheveled-hair'+
pi1 - tou2 si4 'four'
- #7) LASER 雷射 'thunder'+ 'shoot'
lei2 she4
- #8) SONAR 聲納 'sound'+ 'receive'
sheng1 na4
- #9) VITAMIN 維他命 'support'+ '3rd pers'+
wei2 tai ming4 'life'

All of the folk-etymologized loanwords in #5-#9 show a fairly good phonetic similarity to the English model, and all are at least suggestive of the meaning. Only #5 "jelly bean" is the most phonetically accurate transliteration that Mandarin can provide, in #6-#9 there are sacrifices of phonetic similarity made in order to provide the semantic link. In terms of semantics and morphological structure, #5 and #6 are both accurate descriptions of the referent and well-formed Mandarin compounds, while #7 and #8 suffer semantically (what a laser shoots is not "thunder", and sonar does not merely "receive" sound), and the VO structure of #9 is anomalous.

In between the transliteration of loans by seemingly arbitrary characters in #1-#4 and the successful marriage of semantic and phonetic replication in #5-#9 are transliterations using characters that only hint at the meaning, or that mix arbitrary and meaningful characters:

- #10) SHOCK 休克 'rest'+ 'conquer'
xiu1 ke4
- #11) DIOXIN 戴奧辛 'wear'+ 'mysterious'+
dai4 ao4 xin1 'caustic'

- #12) SUNKIST 香 吉 士 'fragrant'+ 'lucky'+
xiang1 ji2 shi4 'scholar'
- #13) FORD 福 特 'lucky'+ 'special'
fu2 te4

In #10 'rest' bears some relation to what your body does when you're in shock, and in #11 'caustic' contributes the idea of a toxic chemical, neither of these is even close to sufficient as a description of the referent, and the additional syllables serve only to replicate the sound of the English signifier. #12 and #13 are examples of a different type of phonetic loan, the transliteration of commercial brand names. These most commonly use characters that have slight connection to the product ('fragrant' in #12 is the only semantically related one), but hve positive or auspicious connotations (characters meaning 'lucky', 'special', 'scholar' etc.) (7). One of the most important jobs that a Taiwan advertising company can do for a foreign client is to choose the proper transliteration characters, ones that link the product to notions of good fortune, beauty and happiness.

3.2. Alphabetic loans

Nearly universal English language training in public schools in Taiwan has not produced a generation of fluent English speakers, but it has produced widespread knowledge of the English version of the Roman alphabet among literate people. This allows increasing borrowing from English by means of graphic loans. Such borrowing creates special problems of adaptations: unlike graphic borrowing between alphabetically written languages, where L2 reading rules can be applied to the graphic form borrowed from L1, Chinese has no reading rules that can be applied to alphabetic script. The creation of reading rules would be not much simpler than teaching everyone English, and would negate any advantage that can be gained from the simplicity of graphic borrowing. Instead, the folk in their wisdom have spontaneously solved the problem through the creation of what I call the Sino-alphabet. (8)

The Sino-alphabet is simply the letters of the Roman alphabet, as used in English, paired with their spoken names, as used in English ([eɪ], [bi], [si] etc.). The exact pronunciation of a given letter depends on the speaker's level of English ability, for instance a speaker with good command of English might pronounce C, H and N as [si], [eɪtʃ] and [ɛn], while most people who speak little or no English will pronounce them [ɕi], [eitɕi], and [ɛn]. The primary feature that distinguishes the Sino-alphabet from its English model

is that it functions non-alphabetically. Letters do not represent phones that are then combined into syllables, they represent invariant units of a syllable or more in length. For instance:

- #14) AIDS "AIDS"
 #15) BIC "BIC" (pens)
 #16) MTV "rock video"
 #17) PVC "polyvinyl chloride"
 #18) X- 光 "X-ray" (X + 'light')
 guang1
 #19) B 型 肝 炎 "hepatitis B" (B +
 xing2 gan1yan2 'type'+ 'hepatitis')
 #20) M 十 六 "M-16" (rifle)
 shi2 liu4

The non-alphabetic, non-combinatory nature of the Sino-alphabet favors borrowing of acronyms like #16-#17 and blends with single-letter tags #18-#20. Forms like #14 and #15 are reanalyzed non-alphabetically, facilitating Sino-alphabetic adaptation, producing the pronunciations [ej aj ti Es] and [pi ai qil]. Borrowings like #14-#20 are increasingly common, and often are the only written form that a lexical item has, not merely stylistic variants or instances of code switching.

4. DISCUSSION

The transliteration characters do not behave like syllabaries in a phonographic writing system. First of all, there is no finite number of them (although some of them are quite common, like 克 ke4 in #1-#2). Second, choice of a transliteration character can be based on semantic principles even when the word is not a full-blown folk etymology, as #10 and #11 show. That is, there is a continuum between fully semanticized transliterations and fully phonetic ones, creating a gray world of fuzzily semanticized loanwords that makes the drawing of a boundary between phonographic and morphographic uses an arbitrary exercise in futility. Third, even when a transliteration character has no semantic relation to the foreign word it transcribes, it cannot be arbitrarily replaced by homophonous characters.

If they follow the vast majority of their fellow characters and represent morphemes, then how can these morphemes have semantic and syntactic properties

entirely different from those morphemes usually represented by these characters? In some cases there is no discrepancy, as in #5. In others, like #7-#11 varying degrees of syntactic and semantic deviation in the use of these morphemes create what Haas (1983) calls "idiomatic" usages, exocentric compounds that are semantically translucent. Yet the total opacity of the loanwords in #1-#4, is due to a complete lack of connection between their meaning and internal structure on the one hand, and the meaning and syntactic functions of their constituent morphemes on the other hand. While the transliteration morphemes in #7-#11 are more or less polysemous with the native morphemes whose graphic form they share, those in #1-#4 are only homophonous, and correspondingly more idiomatic.

The entities represented by transliteration characters are best understood as empty morphemes. These empty morphemes are created by the imposition of prototypical Chinese morphological structure (monosyllabic morphemes, combining to form multisyllabic compound lexical items) on borrowed multisyllabic monomorphemic (or morphologically unanalyzed) signifiers. If a morpheme is a unit that combines three kinds of information-- a phonological value, a semantic value, and syntactic privileges and obligations-- then these morphemes are empty of the latter two, possessing only phonological value.

Being empty is not a permanent state, these morphemes can be and often are "filled". In a phonetic loan, replication of the signifier is accomplished through identification and substitution of phones, with no reference to morphemic identity. In a language like Mandarin with its rampant homophony at the syllabic level, a given syllable can be equally well related to any of a number of native morphemes, and a new criterion comes into play: semantic relatedness to the signified. This represents the filling of the empty morph with a native morph. For example, in #5 JELLY BEAN, the two syllables used to replicate jelly were filled with native morphemes in a combination that is semantically closely related to the signified. In #10 SHOCK, only the first syllable could be filled, (with a more tenuous semantic relationship), the second syllable remains empty. In #1-#4, all remain empty.

It is precisely the previous existence of empty morphemes that has made the formation of the Sino-alphabet possible. The 26 graphs with their 26 different phonological values, empty of any inherent meaning or syntactic restrictions, structurally parallel the empty morphemes used in transliteration. The only difference is that they have no graphic form in common with native, meaningful morphemes, leaving no

possibility for them to be "filled".

5. CONCLUSION

Y-R Chao referred to "...the genius of the Chinese language to read meaning into every syllable" (1968:167). This overwhelming propensity of the language, combined with a writing system that represents morphemes syllable by syllable, creates a prototype of writing as a one-to-one relationship of graph to morpheme to syllable (9). In multisyllabic loanwords, this prototypical relationship cannot exist, yet it is too strong to be ignored. The meaningless syllables of loanwords are treated as empty morphemes, to be filled with whatever morphological or semantic content will bring it closer to the prototype. The empty morpheme strategy is clearly illustrated in the development of the Sino-alphabet, with easy and extensive borrowing of non-alphabetic uses of Roman letters, while borrowing of alphabetic uses is incompatible with the Chinese writing system.

In the classification of writing systems, it is clearly impossible to classify a script as inherently logographic, morphographic, syllabic, or alphabetic. The Sino-alphabet is an example of the Roman alphabet used as morphographic writing. Chinese characters, which function morpho-syllabically in Chinese, can also be fully morphographic (Japanese kanji) or completely syllabic (Japanese man'yogana, the ancestor of kana (Sampson 1985), or Sino-Mongolian (Halliday 1959)). It is a great improvement to say that classification should only apply to the relationship of a given writing system to a given language, as is the case of calling Chinese characters "morphosyllabic" when applied to Mandarin, taking into account the ways that the structure of the language interacts with that of the script. Yet even then discrepancies emerge when phenomena like loanwords are considered. Only when writer/readers are given full credit for understanding and using all the connections between the various elements of their language, both written and spoken, at all levels of structure, can we see principles of writing like the logographic and the alphabetic for what they are-- strategies for relating written to spoken language that are available to any literate person in any written language.

NOTES

(1) The fieldwork portion of the research for this paper was supported by the U.S. Department of Education Fulbright-Hays Research Abroad Program, Grant #022AH60004.

(2) In Mandarin, one glaring exception is the suffix -er, a distinct element graphically and

morphologically, which is incorporated into the preceding syllable. Another class of exceptions is disyllabic monomorphemes, both old loanwords (e.g. 葡萄 pu2 tao2 "grape", 琵琶 pi2 pa2 "balloon lute") and many names for insects and vermin: 蚂蚁 ma3 yi3 "ant", 蝴蝶 hu2 die2 "butterfly", 蜈蚣 wu2 gong1 "centipede", 蚂蟥 ma3 huang2 "leech" etc. This plethora of disyllabic bug names presents a fascinating puzzle, but its very concentration in one corner of the lexicon weakens its status as a counterexample to the one morpheme = one syllable generalization.

(3) Fromkin (1987) presents evidence that readers and writers of English not only use both kinds of information (phonographic and logographic), but that they are stored differently in the brain.

(4) Cheng (1978) explicitly examines different encoding strategies used by Taiwanese speakers when writing morphemes that have no standard written form.

(5) Gao and Liu (1957) give a detailed discussion of Chinese graphic borrowing from Japanese, and its relation to both Western languages and Classical Chinese.

(6) Godwin (1979) finds that in transliteration of foreign names in Hong Kong, there is no specific set of characters that is used, but that a certain structural type of character is favored.

(7) Godwin (ibid.) describes an even more subtle device for introducing connotations into transliterations. In Cantonese transliterations of Western celebrities' names, the radicals 王 "jade", 艹 "grass", and 女 "woman" are frequently used in women's names, but not in men's.

(8) Bauer (1982) and Chan and Kwok (1982, 1986) describe alphabetic forms used in the Cantonese of Hong Kong, without commenting on the theoretical implications for the writing system.

(9) Wang (1981) calls it the "happy fit" between Chinese morphological structure and Chinese characters.

REFERENCES

- Bauer, Robert. 1982. "D for two in Cantonese". Journal of Chinese Linguistics 10:276-80
- Bolinger, Dwight. 1946. "Visual Morphemes." Language 22:333-40
- Chan, Mimi and Helen Kwok. 1982. A Study of Lexical Borrowing From English in Hong Kong Chinese. Centre of Asian Studies, University of Hong Kong.
- , 1986. "The impact of English on Hong Kong Chinese". in W. Viereck, and W-D Bald, eds. English in Contact with Other Languages. Akademiai Kiado: Budapest.

- Chao, Yuen Ren. 1968. A Grammar of Spoken Chinese. University of California Press: Berkeley.
- Cheng, Robert. 1978. "Taiwanese morphemes in search of Chinese characters". Journal of Chinese Linguistics 6:306-13
- DeFrancis, John. 1984. The Chinese Language: Fact and Fantasy. University of Hawaii Press: Honolulu.
- French, M.A. 1976. "Observations on the Chinese script and the classification of writing systems." In W. Haas, ed., Writing Without Letters. Manchester University Press: Manchester. pp. 101-29
- Fromkin, Victoria. 1987. "The lexicon." Language 63:1-22
- Gao, Mingkai and Liu Zhengtan. 1957. Xiandai Hanyu wailaici yanjiu (Research on Foreign Words in Modern Chinese). Wenzhi Gaige Chubanshe: Peking.
- Gelb, I.J. 1963. A Study of Writing. University of Chicago Press: Chicago.
- Godwin, Christopher D. 1979. "Writing foreign terms in Chinese". Journal of Chinese Linguistics 7:246-67
- Haas, William. 1983. "Determining the level of a script." In Florian Coulmas and Konrad Ehlich, eds., Writing in Focus. Mouton: New York.
- Halliday, M.A.K. 1959. The Language of the Chinese Secret History of the Mongols. Basil Blackwell: Oxford.
- Hansell, Mark. 1986. "Lexical Borrowing". UC Berkeley Special Field Exam. ms.
- Haugen, Einar. 1950. "The analysis of linguistic borrowing". Language 26:210-31
- Pulgram, Ernst. 1976. "The typologies of writing systems." In W. Haas, ed., Writing Without Letters. Manchester University Press: Manchester. pp. 1-28
- Sampson, Geoffrey. 1985. Writing Systems. Stanford University Press: Stanford.
- de Saussure, Ferdinand. 1915. Course in General Linguistics. English edition, Translated by Wade Baskin, 1959. McGraw-Hill: New York.
- Wang, William S.-Y. 1981. "Language structure and optimal orthography." In Ovid J.L. Tzeng and Harry Singer, eds., Perception of Print: Reading Research in Experimental Psychology. Erlbaum.
- Weinreich, Uriel. 1968. Languages in Contact: Findings and Problems. Mouton: The Hague.